

Revision memo

‘Do TJ policies cause backlash? Evidence from street name changes in Spain’, submitted to Research & Politics (Manuscript ID RAP-21-0006)

June 8, 2021

Response to Editor

Dear Editor,

We would like to thank you for the opportunity to revise the manuscript. As you will see, we have made thorough revisions, seriously considering each individual point raised by the reviewers. Below we address each of the reviewers in turn.

In general terms, we have XXXX Finally, we have also improved the empirical analyses, addressing all of the reviewers’ points on the analyses. We hope you will find that the manuscript has significantly improved through this process.

Response to Reviewer 1

We would like to thank the reviewer for all the comments. They were all very thoughtful and we found them very helpful in improving the manuscript. We respond to each of them below.

Comment 1: *On a conceptual level, I think the author(s) need to 1) define TJ policies more concretely, and 2) explain how renaming streets fits into the concept of TJ generally. The current definition of TJ offered in the manuscript (bottom of p2) combines truth-seeking, justice, and the politics of memory, but the case itself is exclusively about the politics of memory. The theory/discussion of contributions should reflect this focus on the politics of memory specifically, and highlight how these might be distinct from other types of TJ policies. In short, scope conditions need to be clearer.*

Response: pending

Comment 2: *As it stands, I think the theory needs to be developed. Why do certain individuals react against TJ policies? Also: is asymmetric polarization context-specific to Spain, or a more general phenomenon?*

Response: pending

Comment 3: *On the broader contribution: I think there are two potential interpretations. Is backlash against TJ policies representative of the failure of TJ (ie: has renaming streets actually harmed the TJ process)? Or is backlash a second-order effect of TJ (ie: is successfully removing street signs a TJ victory with unfortunate side effects)?*

Response: pending

Comment 4: *On the empirics, my primary concern is that I can't fully judge whether the DiD is appropriate without more info. The unit of analysis in both analyses is the municipality; the DiD compares municipalities that changed street names to those that left them unchanged. However, the data section provides descriptive stats on street names rather than municipalities. Figures 1 and 2 show the total share of streets names in Spain that are Francoist and the number of individual street name changes, neither of which are relevant to the identification strategy per se. What I want to see:*

- 1) *The total number of municipalities that have/don't have Francoist street names in 2016 (sample size);*
- 2) *The number of municipalities with Francoist street names in 2016 that changed one or more of these street names between 2016 and 2019 (ie: specify how many municipalities*

in the T and C conditions);

- 3) *How many street names were removed on average – and proportional to the total Francoist street names in the municipality – for municipalities that had street names removed (treatment strength);*
- 4) *Some discussion of how municipalities in T/C conditions compare (ie discussion of Table 1 in Appendix, or just descriptive stats);*
- 5) *Whether the trends in Vox/PP/PSOE vote share are parallel leading up to 2016.*

We completely agree with the reviewer that more information is needed to judge the identification strategy, and we think this is a very good point. We have now included more additional information and graphs, both in the main text and in the appendix. In particular:

- We have included a frequency table (table 1) in the main text specifying the number of municipalities included in the DiD sample (i.e. municipalities that had Francoist street names in June 2016) and how many of these were in the treatment (street name removal between June 2016 and December 2018) and in the control groups.
- We have included in the appendix a section called ‘DiD sample and treatment strength’ where we discuss the number of Francoist street name removals in treated municipalities, relative to the number of Francoist street they had at the beginning of the period, and how many Francoist street names they had not been yet removed at the end of the period.
- We include a more thorough discussion of Table 1 in the appendix, discussing the differences between treated and control municipalities in the sample (section A4 in appendix).
- We include in the main text a graph showing the parallel trends for the main parties of the analyses (PP and Vox, following also the recommendation in the next comment), and we include much more detailed tables and analyses in the appendix (including test of the parallel trends in alternative DiD models for PP, Vox and PSOE with data since the early 2000s).

Comment 5: *Also, on parallel trends:*

- 1) *I’d like a plot in the main text showing trends for PP and Vox aren’t diverging in municipalities that removed names (or converging in non-removing municipalities) prior*

to 2016 – ie the potential endogeneity issue discussed at the bottom of p13. This is the key identifying assumption, and presenting tests (such as results from Table 5-6 from appendix, or just trend plots) in the main text is standard.

- 2) Author(s) test the assumption in Appendix Tables 5-6 but not for PSOE. Why not?*
- 3) Vox only comes into existence in 2013, so there is only one election to test the assumption properly. Is the test for parallel trends in Table 5 still valid? Is there anyway to go further back? Some discussion of this is necessary.*

We also agree with the reviewer that showing that there are no diverging trends between control and treated municipalities is key for the credibility of the results. We have now included a plot in the main text showing these parallel trends for PP and Vox, and include much more detailed analyses in the appendix (robustness tests for PP, PSOE, and Vox).

The reviewer is right in saying there is no reason not to include these robustness tests for PSOE. We did not include them in the beginning for reasons of brevity, but we now include a full table repeated all the same robustness tests for the DiD model (including parallel trends) using PSOE as dependent variable. As shown in section A in the appendix, we now include data on elections since March 2000, when PP won its first absolute majority in parliament, thus testing parallel trends across different electoral periods that include right-dominated and left-dominated two-party systems and the emergence of a multi-party system. We also include a reference to these analyses in the main text, even though we cannot discuss their details for reasons of space. Overall, we do not find evidence of significant differences between municipalities in treated and control conditions, in terms of Vox (in 2015), PP (2000-2015) or PSOE (2000-2015). If anything, some models suggests that treated municipalities had shown stronger support for the PSOE in 2011 and 2015, although this result is not robust across all specifications. Moreover, we think that this should not be a concern for the analyses, as both our argument and results focus on a process of asymmetric polarization that mainly affects right-wing voters.

Finally, the reviewer raises a potential concern about the lack of data previous to the treatment for Vox. This is a very relevant point. However, there no easy solution as Vox did not participate in national elections before that year. Yet, we think that given that most electoral supporters of Vox had been PP voters in the past (e.g. ?), testing the parallel trend assumption using PP votes (which holds for every year since 2000) should help to alleviate this potential problem. We have included a footnote in the main text discussing this point.

Comment 6: *On the modeling:*

- 1) *Why do sample sizes in the DiD differ dramatically from outcome to outcome in Table 2? N should not change across outcomes, unless 1) vote shares are much more likely to be missing for Vox or 2) these are different samples – both of which would be a concern.*
- 2) *I'm not sure the cross-sectional analysis adds to the empirics, for the selection reasons the author(s) acknowledge. Maybe a triple difference with a second control category of municipalities that do not have Francoist street names in 2016 would similarly speak to generalizability of effects while retaining the nice identification strategy.*

We agree with the reviewer that different sample sizes are a potential concern. The reason for these differences is that Vox did not participate in all municipalities in 2016, and that is why the sample is slightly smaller in this case. Yet, we agree that we should estimate the models on the sample sample, and even though the results do not change, we now limit the samples in the PP and PSOE models to those municipalities included in the Vox analyses (we explain this in a footnote in the main text, in page **XX**).

Regarding the second point, we thank the reviewer for suggesting this possibility, which we were not aware of. We have extensively look into this method and, even interesting, we think it does not fit this research design and the data we have. As far as we understand diff-in-diff-in-diff (DDD), this method is meant to include a second control category. However, it needs to be linked to the treatment group in some way that excludes the original control group. A classical example is a DiD design which compares the effect of a policy targeted at the elderly on the 55-64 (control) and 65+ (treated) population within a state (see (?, 150–151)), and expands the model into a DDD by introducing as a second control the elderly population (65+) of another state.

In our case, however, there is no way to distinguish the second control (municipalities without Francoist street names in June 2016) from the treated population but excluding the original control (municipalities with streets in June 2016 but that did not change them during 2016–2018). Because of this, the DDD estimate (the triple interaction term $treated \times control1 \times april2019$) in a DDD equation with these three groups fails to estimate.

That said, if we have misunderstood this method and there is indeed a way to apply it to this design, we would be very grateful about some references on how to proceed, because we do think it is very interesting and would greatly improve the analyses.

Comment 7: *Small things:*

- *The author(s) discuss Vox and PSOE, but (how) did PP politicize the issue?*

Response: pending

- *Does the federal government coerce non-changers or reward changers? Might certain municipalities be more/less responsive to incentives?*

Response: pending

- *The author(s) say that the DiD sample is “probably” more rightist than the overall population of municipalities (p9). This can be tested empirically, and would be interesting.*

Response: This is a very good point. We now include a much more thorough comparison of municipalities in and out of the sample (including t-test comparison on electoral shares and logistic regressions) in the appendix (section A4). Also related to more comments by the reviewer above selection bias above, this is a very important point for the analyses. What we find is that indeed the municipalities included in the analyses are slightly more rightist than the rest, particularly when we only compare municipalities within the same region. We also include a reference in the main text to this section in the appendix.

- *I'd like to see a version of the DiD without controls.*

Response: We now include in the table on the main DiD models (table 2 in previous manuscript, table 3 in current one), models without control for each of the parties (Vox, PP, PSOE). Results do not change significantly.

- *Why do the author(s) simulate DiD estimates in Figure 3?*

Response: We use simulations to obtain an estimate of uncertainty of the DiD estimates. We draw 1000 set of coefficients from a multivariate distribution, and based on those draws, we simulate data. We then extract the DiD estimate for each simulation, as $(y_{trt,t1} - y_{ctl,t1}) - (y_{trt,t0} - y_{ctl,t0})$. Figure 3 shows the mean and 0.025 and 0.975 percentiles of these simulations.

- *Are the standard errors clustered on Autonomous Communities?*

Response: We do not use clustered standard errors on Autonomous Communities because we did not think there were reasons to do so: the treatment is not clustered at the level of Autonomous Communities, and we already include fixed effects to account for potentially different levels of regional variation. That said, we agree with the reviewer that it is important to account for potential biases in the SEs, and now include a table in the robustness section of the appendix (section A6) where we show results for the main

models using heteroskedasticity-consistent SEs and SEs clustered at the level of municipalities. Although in some cases the level of significance decreases a bit, and in some cases it increases, all key coefficients retain some level of statistical significance.

- *Figure 2: title says “percentage”, y-axis says “share”.*

Response: We now changed the figure caption so they are coherent.

- *Possible typo on first paragraph of page 14: “dependent variable in continuous form (logged number of street name removals)”. Should this be independent variable? Also: how do the author(s) run a DiD with a continuous treatment?*

That is right, it is a typo. We have now changed it to “the main independent variable”.

About running a DiD with a continuous treatment, we do in the conventional way, including an interaction between the time indicator and the treatment variable (which indicates treatment strength). The only difference is that the DiD estimate should not be interpreted in a binary way (the treatment effect) but as the effect of an increase in the (logged) number of Francoist street name removals on the treated. The other two coefficients are interpreted the same way.

Response to Reviewer 2

We would like to thank the reviewer for all the comments. They were all very thoughtful and we found them very helpful in improving the manuscript. We respond to each of them below.

Comment X: The literature Review and theory could be a bit more focused particularly to provide justification for the backlash idea. At the moment the expectation/hypothesis follows a very general literature review on the effects of transitional justice measures. More focus on the effect of political legacies on political support in the relation to transitional justice would help here as well. I would recommend looking beyond the literature on transitional justice and considering the growing literature on authoritarian legacies and political behavior in new democracies (see for example the Comparative Political Studies Special Issue Neundorff, A. and Pop-Eleches, G., 2020. Dictators and Their Subjects: Authoritarian Attitudinal Effects and Legacies. *Comparative Political Studies*, 53(12), pp.1839-1860.)

Response: pending

Comment X: The initial motivation is a bit confusing by focusing on the US and then switching to Spain. Why the US? Surely taking down symbols of the old regime is a more general phenomenon? Also the US perhaps does not fit the more general paradigm of a post-authoritarian context. And why is Spain a good case to examine the effects of taking down symbols? Are the mechanisms proposed to be at work in Spain generalizable to other contexts? Why do you assume that taking down symbols of the old regime will necessarily increase support for the far-right? Why not extremist parties in general? See Milan Svolik's paper (Svolik, M.W., 2019. Polarization versus democracy. *Journal of Democracy*, 30(3), pp.20-32.) How would what you find be generalized to post-communist contexts? Do you think taking down communist symbols should lead to a revival of some form of left-wing parties?

Response: pending

Comment X: Also how generalizable are the findings even within Spain? Is the backlash specific to areas that might have high latent rightist support? This issue is also linked to concerns about the DiD design.

Response: pending

Comment X: The methodological and empirical sections need to be better organized and more transparent. The relevant information for understanding the DiD design is

scattered throughout the methodology section and the first part of the analysis section. The word count limit is low, but specific information on the measurement levels of each variable would be helpful as well as the model equations. For the robustness checks referred to in the appendix, please provide specific table numbers in the appendix in the main body of the text. It would be clearer to introduce the variables, then run the main analysis, and then refer to the robustness checks in the appendix. As an example the author(s) chose to code the Francoist street name removals as a binary variable. This is an important decision. I would like to see some graphical representation of how many and which municipalities would fall into either category. For the cross-sectional analysis I would be interested in seeing the effects of the Francoist street name removals as an interval measure in relation to the effects of the binary coding. The authors say they have done the additional analysis using an interval measure (logged number of changes) but only in the methods section (page 6) and also it is not very clear which table they refer to in the appendix. It would be logical to mention this robustness check after the discussion of the results in Table 1 of the main body. It is not clear which table shows the cross-sectional analysis with the logged number of removals.

Response: pending

Comment X: I would also generally be interested in understanding more about what distinguishes the municipalities that retained Francoist street names until the periods analyzed from those that removed the names before 2001. Perhaps a map of Spain showing where are the municipalities included in the sample are located would be nice. What is their history? Are they former Francoist strongholds? If so the effects of removal of symbols would be rather inflated: that is if we could go back in time and look at the earlier effects of the removals of street names would they have caused an equally sized backlash? Maybe not if the removals were done in areas of Spain that were more anti-Franco.

Response: pending

Comment X: For the first cross-sectional analysis: what is the point of the model (Table 1) on the full sample if the majority of municipalities had no Francoist street names to remove? In other words these municipalities are not eligible for “treatment” anyhow and more likely to be fundamentally different from the municipalities that still retained Francoist street names in 2001 or later. For the cross-sectional analysis I would be interested to see the effects of the interval/continuous measure of Francoist street name removals. I would also like to systematically see maybe in the appendix side by side

the regression results using varied periods in which the Francoist street name removals are considered (e.g. 2001-2018, 2011-2018, etc). The main point is that we have some systematic change in the samples so for example same end point to the period but sliding starting points. I have also noticed that the samples vary according to whether all the included municipalities had some Francoist street names in the starting year of the period considered for the independent variable. It is logical obviously to look at the municipalities that had Francoist street names at the starting but again consistency would help (Table 4 in the appendix states the sample only includes municipalities that had street names in June 2011 but the period considered for removal is December 2010-2018).

Response: pending

Comment X: The application of the difference-in-difference design could be much clearer. The word count limit is quite short but briefly writing out the model (Also clearly specify which in the models coefficients capture the treatment effect) and perhaps explaining the design explicitly in terms of treatment and control would help (additional details could go in the appendix). Noting details like the model type would help too (I am assuming OLS based on the dependent variable). Overall it seems like the design should be viable and defensible although the generalizability of the findings might be more limited than currently stated. But I would like to see more explicit consideration of DiD design. To this end it would help write out the model and in the appendix maybe explain the more design more systematically in terms of treatment and control and expected outcomes. Telling us the mean values on the outcome for these municipalities versus the “treatment” municipalities as well as the levels of Francoist street names.

Response: pending

Comment X: Next the author(s) briefly mention the issue of “selection bias” and how the “control” group municipalities could be more right-wing. I would like to know more about all this and how the design is still justifiable. Let us assume for example that the municipalities in the control group that have retained Francoist street names and resisted changing them did so because they are right-wing strong holds, how would this characteristic affect support for Vox? Is it possible that the support for Vox was established earlier before the period analyzed and reached a ceiling? So maybe we do not have much change/variation in Vox support to explain among the control group because the support appeared earlier and remained stable? I am also puzzled whether the level of name changes is the only likely trigger for the increase

in Vox support. If there is so much potential for far-right support, who was in control of these municipalities and who drove through the street name removals? Was it left-wing local governments? Are we seeing perhaps the effects of municipal level polarization driven by the national discourse? The author(s) note that two events coincided: the removal of street names and Vox nationalist campaign. Could both be responses to overall political polarization? It would be also interesting to have more discussion about the finding that support for the more relatively more centrist PP declined. Again see debates related to Svoboda 2020. Some of these issues might partially be addressed through interpretation of Table 1 in the appendix.

Response: pending

Comment X: Also on a more detailed technical note what about first differencing the DiD model? Then you could better leverage the variation in the treatment effect. See discussion of Card 1992 in the section on DiDs in Angrist, J.D. and Pischke, J.S., 2008. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Response: pending

Comment X: I would also like a clearer and more explicit discussion of the parallel trends assumption in the appendix. It would help to see some visual representation of party support over time in the municipalities included in the DiD analysis. I would expect to see that that difference (could be no difference) in party support for the right-wing parties (PP, Vox) between in the "control" municipalities (those that had Francoist street names and did not change them) and those (that removed Francoist street names) remained stable before the "treatment period." For the support for Vox it appears that there is only one time point (2015 election) before the pre- and post-treatment time periods. Really to establish a pre-treatment parallel trend more than one time point is needed ideally. Maybe the PP vote could be a proxy. We should expect the PP vote relative to the Vox vote to be stable in the control group and stable difference between PP votes in the treatment vs control going back in time.

Response: pending

Comment X: As to the conclusions and implications: why should the removal of old regime symbols just increase in support for the far-right? If the idea of backlash is to travel, then surely we would expect in Eastern Europe an increase in support for the left if communist symbols are removed?

Response: pending

Again, we very much appreciate blah blah